5110–5116 Nucleic Acids Research, 1993, Vol. 21, No. 22

Reconstitution of mammalian excision repair activity with mutant cell-free extracts and XPAC and ERCC1 proteins expressed in *Escherichia coli*

Chi-Hyun Park and Aziz Sancar*
Department of Biochemistry and Biophysics, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA

Received July 30, 1993; Revised and Accepted September 16, 1993

ABSTRACT

Nucleotide excision repair in humans involves the coordinated actions of 8 – 10 proteins. To understand the roles of each of these proteins in excision it is necessary to develop an *in vitro* excision repair system reconstituted entirely from purified proteins. Towards this goal we have expressed in *E.coli* two of the 8 genes known to be essential for the excision reaction. XPAC and ERCC1 were expressed as fusion proteins with the *Escherichia coli* maltose binding protein (MBP) and purified to > 80% homogeneity by affinity chromatography. The purified proteins either as fusions or after cleavage from the MBP were able to complement the CFE of cells with mutations in the corresponding genes in an excision assay with thymine dimer containing substrate.

INTRODUCTION

Nucleotide excision repair is the major molecular mechanism for removing modified bases from DNA. Genetic studies with the human hereditary disease xeroderma pigmentosum (XP) have identified 7 genes (*XPAC–XPGC*) involved in the incision reaction (1). In addition a number of UV-sensitive rodent cell lines have been isolated (2) and at least one of these is defective in a gene, *ERCC1*, not represented by the XP complementation groups. Thus, to reconstitute human excision repair *in vitro* at least 8 proteins are required (3,4).

An *in vitro* excision repair system (5,6) is available for complementation studies between cell-free extract (CFE) from mutant cell lines (4,5,7) and the system has been used in complementation assay for purifying repair proteins (8). Since repair proteins are, as a rule, present in the cell in small quantities, purifying all of the proteins thought to be required for excision from their natural source might not be practical. Since most of these genes have now been cloned, it is reasonable to attempt to express the proteins in heterologous systems and purify them with the ultimate goal of reconstituting human excision repair *in vitro*. We have initiated such studies using two genes whose products are thought to be involved in either damage recognition, *XPAC* (8,9) or in nuclease function, *ERCC1* (10). Here, we describe the expression of these genes in *E.coli* in soluble form as fusion proteins with the *E.coli* maltose binding protein, two step purification of the proteins, and functional complementation with the fusion proteins with or without enzymatic digestion to release the human protein. These results are encouraging especially in light of recent findings (4) suggesting that native ERCC1 may not be purified to homogeneity by a functional assay because it is in a tight complex with ERCC4 in vivo and that the formation of such a complex requires the co-synthesis of the two proteins.

MATERIALS AND METHODS

Materials

Restriction enzymes were obtained from GIBCO-BRL and New England Biolabs. Taq DNA polymerase was purchased from Promega. The expression vector pMAL-c2, amylose resin, and Factor Xa protease were obtained from New England Biolabs. Heparin agarose, blue agarose, and DNA cellulose were purchased from Sigma. Oligonucleotides were synthesized at the Lineberger Comprehensive Cancer Center (University of North Carolina) using an Applied Biotechnology Systems DNA synthesizer. Human cDNA libraries were obtained from Clontech.

Cloning of *XPAC* cDNA

Tanaka *et al.* (11) reported the cloning and sequencing of *XPAC* but the clone has not been made available for general use. Therefore, we isolated cDNA clones carrying *XPAC* from commercial cDNA libraries. The λgt11 and λgt10 (5′-stretch) HeLa cell cDNA libraries from Clontech Laboratories, Inc. (Palo Alto, CA) were screened by plaque hybridization. The initial screening of the λgt11 library was performed with a ^{32}P-labeled 36-mer complementary to bp (679–714) of the published *XPAC* cDNA sequence (11). After screening about 1.5×10^6 plaques, one positive clone was identified (pCH1). Sequencing the insert by the dideoxy method revealed that the clone contained a 970 bp segment corresponding to bp(674–1643) of *XPAC* cDNA and

* To whom correspondence should be addressed

597 bp extra DNA (presumably an intron) at the 5′ end of the coding sequence.

Using the insert in pCH1 as a template, a probe was prepared by random primer labeling method (12) and used to screen the λgt10(5′-stretch) library. Five positive clones were identified among the 10^6 plaques screened. One clone (pCH2) was selected for further analysis. Sequencing revealed that the cDNA in this clone contained a 1028 bp-insert corresponding to bp (259−1286) of *XPAC* cDNA with 22 bp of unrelated sequence at the 5′ end of the insert. Another probe was prepared by random primer labeling method using the insert in pCH2 as template, and the λgt10(5′-stretch) library was screened again. Of about 2×10^6 plaques screened, four positive clones were identified; and one of these, pCH3, was selected for further analysis. The insert in this clone contained bp(33−726) of *XPAC* cDNA with 63 bp of unrelated sequence at the 5′ end.

A plasmid carrying the entire *XPAC* cDNA was constructed by ligating the appropriate restriction fragment from pCH2 and pCH3 with a synthetic oligomer encoding the 5′terminal 33 bases. The synthetic oligomer consisted of two complementary oligonucleotides (a 55-mer and a 57-mer) which were synthesized to contain bp(1−53) of *XPAC* cDNA plus 5 additional nucleotides to generate a SacI site at the 5′ end. The two oligomers were annealed and ligated into pCHA302 which contained the rest of the gene to generate pCHA801 which carries the entire *XPAC* cDNA.

Construction of pMAL-XPAC plasmid

An 857 bp fragment of pCHA801 carrying the entire *XPAC* (846 bp) was amplified by PCR. The sense primer was a 21-mer corresponding to nt(1−21) of *XPAC* and the antisense primer was a 25-mer corresponding to nt(833−857) (but with a T to G replacement at position 847 to generate an XbaI site) of the *XPAC* cDNA sequence. The PCR product was extracted with chloroform/isoamyl alcohol (24:1), digested with XbaI, and ligated with the pMAL-c2 vector that had been digested with XmnI and XbaI. The resulting plasmid, pMAL-XPAC contains in-frame fusions of *malE-XPAC-lacZα* as shown in Figure 1.

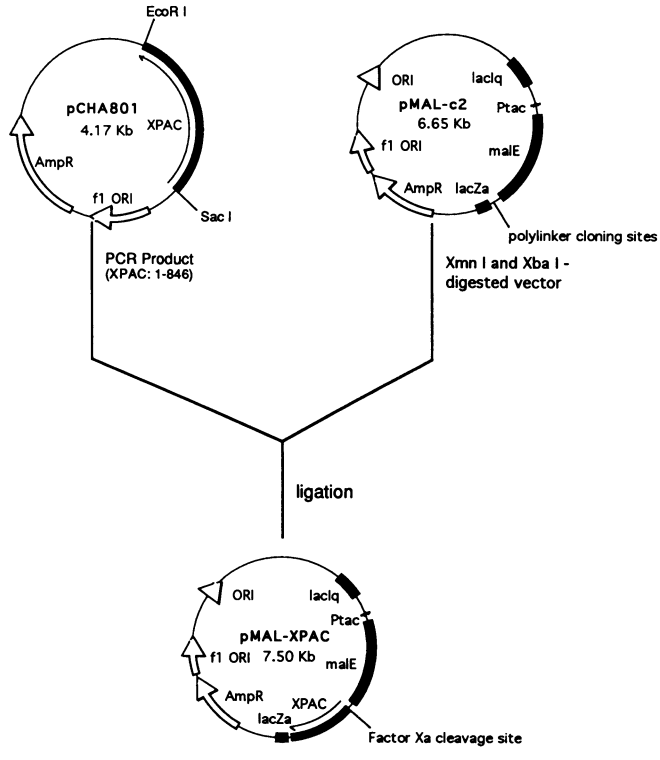

Figure 1. Construction of pMAL-XPAC. The 846-bp XPAC cDNA fragment was generated by XbaI digestion of an 857-bp PCR product of pCHA 801 and inserted into the pMAL-c2 expression vector that had been digested with XmnI and XbaI.

Overproduction and purification of MBP-XPAC

E.coli DR153(*recA*$^-$ *uvrB*$^-$) (13) was found to be optimal among the several *E.coli* strains tested as a host for pMAL-XPAC. Overnight cultures of DR153/pMAL-XPAC were grown in LB + 100 μg/ml ampicillin. Twenty ml of the overnight culture were used to inoculate 2 liters of LB containing 0.2% glucose and 100 μg/ml ampicillin. The cells were grown at 37°C to A600=0.8 at which time IPTG was added to 0.3 mM and induction was carried out for 4 hrs. The cells were collected by centrifugation, rinsed with 40 ml of Rinse Buffer (10 mM Tris−HCl, pH 7.4, 10 mM NaCl, 1 mM EDTA) and resuspended in 40 ml of Buffer A (20 mM Tris−HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA, 10 mM β-mercaptoethanol) containing 5 μg/ml pepstatin A and 5 μg/ml leupeptin. The cells were frozen in dry-ice ethanol and stored at −80°C.

All of the purification steps were carried out at 0−4°C and aliquots of the chromatographic fractions were analyzed by SDS-PAGE to locate the fusion protein. Frozen cells were thawed at 4°C and sonicated 10×10 sec with a Branson model W185 sonicator set at maximal output for the small tip. The lysate was clarified by centrifugation at 9000×g for 30 min followed by centrifugation at 100,000×g for 2 hrs. The supernatant was diluted 1:5 with buffer A and loaded onto a 20 ml amylose column equilibrated with the same buffer, at a flow rate of 1 ml/min. The column was washed with 100 ml of Buffer A and the bound protein was eluted with Buffer A + 10 mM maltose. Four ml fractions were collected and analyzed by SDS-PAGE and Coomassie blue staining.

The fractions containing MBP-XPAC were pooled (16 ml), diluted 1:2 with Buffer B (25 mM HEPES-KOH, pH 7.6, 1 mM EDTA, 10 mM β-mercaptoethanol, 10% (v/v) glycerol) and loaded onto an 8 ml heparin agarose column equilibrated with Buffer B + 0.1 M KCl, at a flow rate of 10 ml/hr. The column was washed with 56 ml of the same buffer and the bound proteins were eluted with Buffer B + 1.0 M KCl. Three ml fractions were collected and analyzed by SDS-PAGE. The fractions containing the fusion protein were combined (9 ml). One ml of this was dialyzed against storage buffer (25 mM HEPES-KOH, pH 7.9, 100 mM KCl, 12 mM $MgCl_2$, 0.5 mM EDTA, 2 mM DTT, 16.5% (v/v) glycerol), divided into 50 μl aliquots, frozen in dry ice-ethanol, and stored at −80°C.

To purify XPAC in free form, the remaining 8 ml (1.2 mg) of the fusion protein was dialyzed against Buffer B + 0.1 M KCl, treated with Factor Xa at 1:100 (w/w) ratio of protease to fusion protein for 12 hrs at 4°C and the extent of cleavage was monitored by SDS-PAGE. The cleavage reaction mixture was then loaded onto an 8 ml DNA cellulose column equilibrated with Buffer B + 0.1 M KCl, at a flow rate of 10 ml/hr. The column was washed with 5 column volumes of the same buffer, and XPAC was eluted with Buffer B + 1.0 M KCl. Three ml fractions were collected, XPAC was located by SDS-PAGE, and the fractions containing the protein were combined and dialyzed

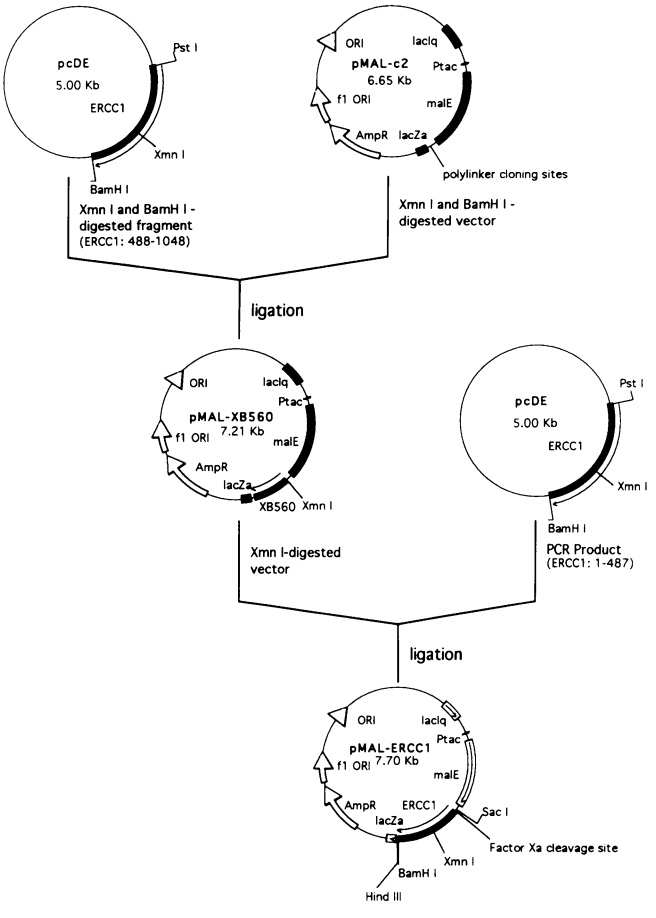

Figure 2. Construction of pMAL-ERCC1. The plasmid was constructed in two steps as shown. First, the 3′ terminal bp(488−1048) and then the PCR-amplified 5′-terminal bp(1−487) half was inserted into pMAL-c2.

against storage buffer. The protein was divided into 50 μl aliquots, frozen in dry ice-ethanol, and stored at −80°C.

Construction of pMAL-ERCC1 plasmid

The plasmid, pcDE carrying the *ERCC1* cDNA (14,15) was kindly provided by Dr J.H.J.Hoeijmakers (Erasmus University, Rotterdam, The Netherlands). A plasmid for expressing a MBP-ERCC1 fusion protein was constructed as follows (Fig. 2). A 560 bp fragment corresponding to bp(488−1048) of the *ERCC1* cDNA was excised from pcDE by XmnI and BamHI digestion and cloned into pMAL-c2 (pMAL-XB560). The bp(1−487) of the *ERCC1* cDNA was amplified by PCR and ligated into the XmnI site of pMAL-XB560 to obtain pMAL-ERCC1.

Overproduction and purification of MBP-ERCC1

The MBP-ERCC1 protein was overproduced, and purified through the amylose and heparin agarose columns by a procedure essentially identical to that used for the MBP-XPAC protein, except that the starting material was a 12-liter culture of DH5αFlacI^q/pMAL-ERCC1 induced for 12 hrs with 0.3 mM IPTG. The peak fractions from the heparin agarose column were combined (12 ml) and one ml of the combined fraction was dialyzed against storage buffer. The protein was stored in 50 μl aliquot at −80°C.

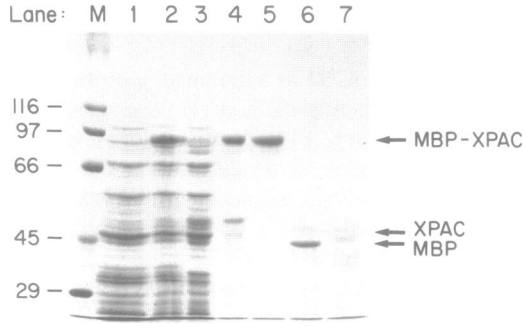

Figure 3. Purification of MBP-XPAC and XPAC. Samples from each of the purification steps were separated on 10% SDS-PAGE and the gel was stained with Coomassie blue. Lanes: M, molecular weight markers with the mass given in kDa; 1, uninduced DR153/pMAL-XPAC cells (250 μl); 2, induced cells (250 μl); 3, cell-free extract (32 μg); 4, amylose column peak fraction (4 μg); 5, heparin agarose peak fraction (5 μg); 6, heparin agarose peak fraction digested with Factor Xa (4 μg); 7, DNA cellulose peak fraction (3 μg).

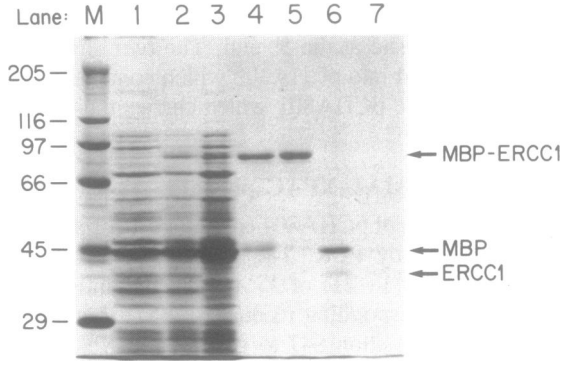

Figure 4. Purification of MBP-ERCC1 and ERCC1. The proteins were separated on a 10% SDS-PAGE and stained with Coomassie blue. Lanes: M, molecular weight markers; 1, uninduced DH5αFlacI^q/pMAL-ERCC1 (500 μl); 2, induced cells (250 μl); 3, cell-free extract (65 μg); 4, amylose column peak fraction (5 μg); 5, heparin agarose peak fraction (5 μg); 6, protein from previous step digested with Factor Xa (5 μg); 7, blue agarose fraction (1 μg).

The remaining protein was diaylzed against Buffer B + 0.1 M KCl, treated with Factor Xa at 1:100 (w/w) ratio for 18 hrs at 4°C. The sample was then loaded onto a 12 ml blue agarose column equilibrated with Buffer B + 0.1 M KCl, at a flow rate of 10 ml/hr. The column was washed with 60 ml of the same buffer and then a 50 ml gradient of 0.1 M to 2.0 M KCl in Buffer B was applied. Three ml fractions were collected, ERCC1 was located by SDS-PAGE and fractions containing the ERCC1 protein were combined (15 ml) and dialyzed against storage buffer. The protein was stored in 50 μl aliquots at −80°C.

Complementation of CFE from XP-A and ERCC1 mutants with purified proteins

CFE from XP-A lymphoblasts (GM02345B), and CHO-ERCC1 mutants (UV20) (16) were prepared as described previously (17) and stored at −80°C until use. The substrate was pUNC1991-4, a synthetic plasmid containing four thymine dimers (T< >T) at predetermined sites and a ^{32}P label at the 11th phosphodiester bond 5′ to each thymine dimer (18).

An *in vitro* excision assay was employed to test the activity of purified proteins. This assay measures the release of 27–29 nt-long oligomers from pUNC1991-4 by the human excision nuclease (18,19,20). Extracts from both XP-A and ERCC1 mutant cell lines are completely defective in this activity (4). The excision complementation assay was conducted as described previously (4) with the exception that complementation was done either by mixing CFE from two mutant cell lines or a mutant CFE supplemented with the missing protein purified from *E.coli*. The reaction mixture (50 μl) contained 40 mM HEPES, pH 7.9, 70 mM KCl, 8 mM $MgCl_2$, 2 mM ATP, 20 μM each dNTPs, 1.2 mM dithiothreitol, 0.3 mM EDTA, 10.2% glycerol, 100 μg bovine serum albumin, 50 μg phosphoenolpyruvate, 5 units pyruvate kinase, 366 ng pUC18 as carrier DNA, 200 ng substrate (~ 200,000 cpm), and 100 μg of CFE plus the indicated amounts of purified proteins in complementation assays. The reaction mixture was incubated at 30°C for 90 min. Following deproteinization by treatment with proteinase K and phenol extraction (5,6), the DNA was precipitated with ethanol, resuspended in formamide/dye and separated on 10% polyacrylamide sequencing gels.

RESULTS

Purification of MBP-XPAC and XPAC proteins

The XP-A protein (XPAC) has been purified from calf thymus (8,21), and from *E.coli* cells expressing the cloned cDNA (22). Although the natural source is ideal for obtaining the protein, the scarcity of it in tissues makes it difficult to obtain quantities sufficient for extensive biochemical and physical characterization. Similarly, the *E.coli* expression system of Masutani *et al.* (22) yields XPAC in inclusion bodies from which active protein must be obtained by denaturation/renaturation. As an alternative we expressed XPAC as a fusion protein with MBP with the aim of obtaining a soluble protein that can be purified by affinity chromatography. Figure 3 and Table 1 show that the fusion protein is greatly overproduced and that much of the overproduced protein is soluble. As a consequence 1.2 mg of protein with 90% purity is obtained by a two-column purification procedure from a 2-liter culture. The majority of the contaminants are degradation products of the fusion protein or the MBP itself, because they are not seen with CFE from *E.coli* without the fusion protein. After the second column the protein is free of nucleases (data not shown) or other *E.coli* proteins.

We considered the possibility that the MBP 'tail' of the fusion protein may interfere with the interactions of XPAC with other human proteins which are necessary to bring about the excision nuclease activity, and thus decided to obtain XPAC protein free of MBP. The fusion protein contains a Factor Xa cleavage site at the junction of its two components. Treatment of the protein with Factor Xa did cleave it to generate MBP and XPAC of the predicted sizes. However, contrary to our expectation the two proteins were not in stoichiometric amounts, the ratio of MBP:XPAC being about 5:1 (Figure 3, lane 6). This is due, in part, to the degradation of XPAC by Factor Xa and, in part, to the precipitation of free XPAC. Nevertheless, the remaining full-length XPAC appears to have native conformation as it binds specifically to a single-stranded DNA cellulose column and can be separated from most of the MBP by passing the sample through this column (Figure 3, lane 7). As it turns out both the free and the fusion forms of XPAC are functional (see below).

Purification of MBP-ERCC1 and ERCC1 proteins

The ERCC1 protein has not been purified before, either from its natural source or by expression in a heterologous system. First, we attempted to express ERCC1 without fusing to another protein, either from a T5 promoter or a T7 promoter, using the appropriate *E.coli* host/vector systems. In the former case the protein was not overproduced, and in the latter case it was overproduced but was insoluble. Therefore, we made the MBP-ERCC1 construct with the aim of obtaining a soluble and functional protein.

Table 1. Purification of MBP-XPAC and XPAC proteins

Fractions	Total volume ml	Protein mg	MBP-XPAC (XPAC) %	mg	Yield %
1. Cell free extract	200	206.25	2.4	4.95	100
2. Amylose	16	2.22	60.4	1.34	27.1
3. Heparinagarose	9	1.36	90.0	1.22	24.6
4. Factor Xa digestion[a]	(8)	(1.21)	(16.9)	(0.20)	(4.1)
5. DNA-cellulose	(6)	(0.43)	(25.8)	(0.06)	(1.2)

[a] Only 8 ml of the previous fraction was processed after this step and the yield is expressed relative to this amount.

Table 2. Purification of MBP-ERCC1 and ERCC1 proteins

Fractions	Total volume ml	Protein mg	MBP-ERCC1 (ERCC1) %	mg	Yield %
1. Cell free extract	360	894.32	1.0	8.94	100
2. Amylose	12	2.44	50.8	1.24	13.9
3. Heparin agarose	12	1.09	81.8	0.86	9.6
4. Factor Xa digestion[a]	(11)	(1.00)	(20.3)	(0.20)	(2.2)
5. Blue Sepharose	(15)	(0.30)	(32.5)	(0.10)	(1.1)

[a] Only 11 ml of the previous fraction was processed after this step and the yield is expressed relative to the fraction that was processed.

Figure 4 and Table 2 show that upon induction of *E.coli* cells carrying the pMAL-ERCC1 plasmid the fusion protein is overproduced in soluble form, and constitutes 1% of total soluble cellular proteins. Although a significant fraction of the fusion protein does not bind to the amylose affinity column for unknown reasons, sufficient amounts bind such that after the amylose column the protein is obtained at about 14% yield and 51% purity, with the major contaminant being the MBP (Figure 4, lane 4). The MBP is mostly removed by chromatography on heparin agarose to yield the MBP-ERCC1 fusion protein at about 82% purity and 10% yield. The major contaminants were degradation products (as evidence by the facts that control runs with CFE from strains carrying the pMAL-c2 vector yielded only the MBP) and thus the protein at this stage of purification was free of non-specific nucleases or DNA binding proteins from *E.coli*. To obtain ERCC1 free of MBP the fusion protein was cleaved with Factor Xa. Again, ERCC1 was released in substoichiometric amounts relative to MBP, possibly because of partial degradation of ERCC1 by Factor Xa. Nevertheless, as in the case of XPAC, the released ERCC1 could be purified by chromatography. The released ERCC1, in contrast to XPAC, did not bind to the DNA cellulose column; however, it did bind to the blue agarose column and was enriched relative to MBP after chromatography on this resin (Figure 4, lane 7). Thus, we were able to isolate free ERCC1, albeit at considerably lower yields compared to the fusion protein.

Complementation with MBP-XPAC and MBP-ERCC1

The purified XPAC and ERCC1 proteins were tested both as fusion proteins and as full-length isolates for complementation of CFE of human XP-A and CHO-ERCC1 mutant cell lines. Fig. 5A shows that full complemenation is achieved with 100 ng of MBP-XPAC (lane 4) and substantial complementation is obtained with 200 ng of MBP-ERCC1 (lane 7). These correspond to about 25−50 nM proteins which are realistic concentrations for repair proteins. Thus, it appears that the 'MBP tail' of the fusion proteins does not interfere with the repair proteins performing their functions. In contrast, the 'free' XPAC and ERCC1 proteins produced by proteolytic cleavage were less active. About 1 μg of each was required to achieve even partial

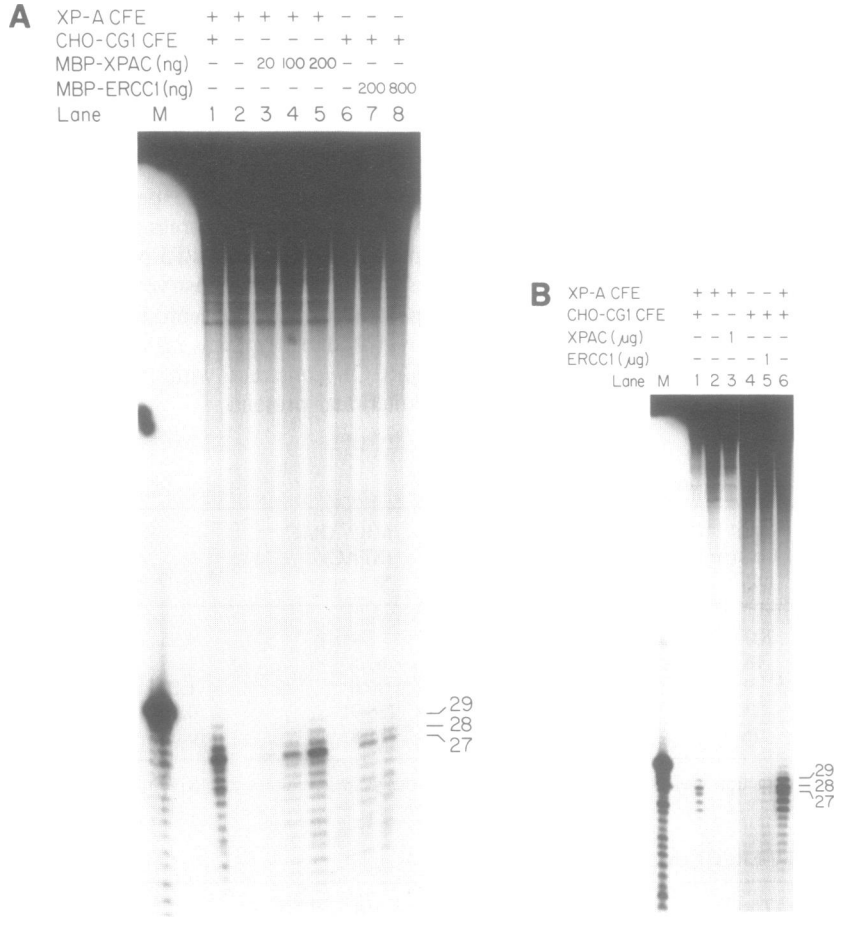

Figure 5. Complementation of CFE from XP-A and ERCC1 (CHO-complementation group 1 = CHO-CG1) mutants with proteins made in *E.coli*. ^{32}P-labeled pUNC1991-4 (T< >T) was incubated with the indicated CFE and purified protein combinations and analyzed on 10% polyacrylamide sequencing gels. The positions of the major excision products (27, 28, 29) are indicated. **A**. Complementation with fusion proteins. M, 30-mer marker ladder; lane 1, XPA-CFE (50 μg) + ERCC1-CFE (50 μg); lane 2, XPA-CFE (100 μg); lanes 3, 4 and 5, XPA-CFE (100 μg) plus 20 ng, 100 ng, and 200 ng of MBP-XPAC, respectively; lane 6, ERCC1-CFE (100 μg). The low molecular weight species seen in this lane are generated by non-specific nucleases in the cell-free extract. Lanes 7 and 8, ERCC1-CFE (100 μg) plus 200 ng and 800 ng of MBP-ERCC1, respectively. **B**. Complementation with XPAC and ERCC1 cleaved from the fusion proteins. M, 30-mer marker ladder; lanes 1 and 6, XPA-CFE (50 μg) plus ERCC1-CFE (50 μg); lane 2, XPA-CFE (100 μg); lane 3, XPA-CFE (100 μg) + XPAC (1 μg); lane 4, ERCC1-CFE (100 μg); lane 5, ERCC1-CFE (100 μg) + ERCC1 (1 μg).

complementation (Figure 5B, lanes 3 and 5). This low specific activity is partly due to the fact that the 'free' XPAC and ERCC1 preparations contain degradation products and thus only a fraction of the 1 μg is full length protein. However, even when correction is made for the degradation products the 'free' XPAC and ERCC1 are still less active than the fusion proteins, presumably because of lower stability of the non-fusion proteins.

DISCUSSION

Genetic and biochemical data indicate that 8−10 proteins are involved in carrying out excision repair in humans (see 3,4). The reconstitution of the human excision nuclease *in vitro* in a defined system would require the purification of all of these proteins. In principle this can be accomplished by purifying all these proteins from their natural source, by purifying the proteins from heterologous systems using cloned genes, or by a combination of both methods. The advantage of the natural source is that the proteins of interest most likely exist in native form and have undergone the necessary posttranslational modifications which are required for certain proteins to attain their functional state. The main disadvantage of the natural source is that repair proteins, as a rule, are not abundant proteins and thus extensive purification is required to obtain sufficient quantities of subunits for enzymological studies. The advantage of the heterologous sources is that large quantities can be produced by using the appropriate host/vector systems, thus making it possible to obtain pure proteins with few purification steps. The disadvantages are that human proteins expressed in heterologous systems are often insoluble, bacteria do not carry out most of the posttranslational modifications made by human cells, and finally the entire set of human excision repair genes is not yet available. Thus, both approaches might be needed to achieve the ultimate goal of reconstituting excision repair in vitro. Here we have described the overproduction and purification of two human repair proteins in *E.coli* and have shown that these proteins are functional in a complementation assay.

The XPAC protein has been purified by both methods mentioned above. Robins *et al.* (8) and Eker *et al.* (21) purified the protein from calf thymus by using either the *in vitro* repair synthesis assay (8) or the *in vivo* unscheduled DNA synthesis assay following microinjection (21) to follow the protein through the purification steps. The purified protein was used to show that XPAC may be the damage recognition subunit of the human excinuclease. Although some reservation regarding its specificity for damaged DNA has been expressed (21) the recent demonstration of high affinity binding to (6-4) photoproducts by RAD14 which is the yeast homologue of XPAC (9) has provided additional evidence for the role of XPAC in damage recognition. In addition to the purification from calf thymus, XPAC has also been expressed in *E.coli* by Masutani *et al.* (22) in an insoluble form, renatured and shown to be active in an *in vitro* repair synthesis complementation assay. The results presented in this paper show that XPAC can be expressed in *E.coli* as a soluble fusion protein with MBP and that the fusion protein, which can be purified in large quantities appears to be as functional as the protein from its natural source. Furthermore, the fusion protein binds tightly to the amylose resin to produce an XPAC affinity column which can potentially be used to identify and isolate other proteins which interact with XPAC.

In contrast to the extensive studies that have been carried out on the XPAC gene and protein, very little is known about the biochemical properties of the ERCC1 protein even though this was the first human excision repair gene cloned (14,23) and it is known to be essential for excision repair (4,24). Sequence comparison revealed that ERCC1 has significant homology to the *S.cerevisiae* excision repair protein RAD10 (15). RAD10 is known to stimulate annealing of complementary DNA strands (25) and to be in a complex with another excision repair protein, RAD1 (26,27), which has a single-strand specific nuclease activity (10). Although no direct evidence exists as yet for complex formation between ERCC1 and other excision repair proteins, *in vitro* complementation studies by Reardon *et al.* (4) have revealed some unexpected complementation patterns involving ERCC1 which have been interpreted in terms of a heterodimer formation. It was found that CFE of *ERCC1* mutant was not complemented by CFE from *ERCC4* mutant. Furthermore, CFE from XP-F cell lines was only marginally complemented by CFE from *ERCC1* mutants and was not complemented at all by CFE of *ERCC4* mutants. It was proposed that *ERCC4* and *XPFC* were identical and that this gene (*XPFC/ERCC4*) was the human homolog of RAD1 and its product made a complex with ERCC1. It was further proposed that the *in vitro* concentration or conditions of ERCC1 and ERCC4/XPFC were not conducive for heterodimer formation (4, see also 28). A corrollary of this conclusion would be that it would not be possible to complement CFE of either ERCC1 or ERCC4/XPFC mutants with the purified proteins. Therefore, the complementation we achieved with MBP-ERCC1 was somewhat surprising.

The most likely explanation for the discrepancy between the complementation experiments with CFE and with the purified protein is that the concentrations of ERCC1 and ERCC4/XPFC are sufficiently high *in vivo* to make complex formation possible but are diluted in the CFE and thus unable to form the heterodimer. If that is the case, then, by increasing the concentration of either or both components of the complex it should be possible to shift the equilibrium towards heterodimer formation. The results presented in this paper lend strong support to the proposed model. With 200 ng of ERCC1 protein we achieved substantial complementation of excision activity of CFE from the *ERCC1* mutant cell line. Since in the complementation assays only 50−100 μg of each CFE is used and, since repair proteins are typically present at a level of about $10^{-4}-10^{-5}$ of total cellular proteins (see 8), the addition of 200 ng of ERCC1 to 100 μg of CFE means increasing the concentration of this protein 10^1-10^2-fold compared to that which can be supplemented by CFE from *ERCC4* mutants. Apparently such an increase is sufficient to shift the ERCC1 + ERCC4 <----> ERCC1/ERCC4 equilibrium towards the heterodimer form.

In conclusion, in this study we have shown that expression of human excision repair genes in *E.coli* as fusion proteins is a promising approach to isolate the human repair proteins in large quantities and to eventually reconstitute the entire human exicison nuclease in a defined system.

ACKNOWLEDGEMENTS

This work was supported by the NIH grant GM32833 and by a grant from the Human Frontier Science Program. We wish to express our gratitude to Dr J.H.J.Hoeijamakers for supplying us the *ERCC1* cDNA clone which made this work possible. We thank Drs. J.T.Reardon and C.P.Selby for critically reading the manuscript.

REFERENCES

1. Cleaver, J.E., and Kraemer, K.H. (1989) In the Metabolic Basis of Inherited Disease, Vol. II. Sciver, C.R., Beaudet, A.L., Sly, W.S., and Valle, D. (eds.) McGraw-Hill, New York, Vol. II, pp. 2949–2971.
2. Thompson, L.H., Busch, D.B., Brookman, K., Mooney, C.L., and Glaser, D.A. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3734–3737.
3. Hoeijmakers, J.H.J. (1993) *Trends in Genetics* **9**, 211–217.
4. Reardon, J.T., Thompson, L.H., and Sancar, A. (1993) Cold Spring Harbor Symp. Quant. Biol. **58**, in press.
5. Wood, R.D., Robins, P., and Lindahl, T. (1988) *Cell* **53**, 97–106.
6. Sibghat-Ullah, Husain, I., Carlton, W., and Sancar, A. (1989) *Nucleic Acids Res.* **17**, 4472–4484.
7. Jones, C.J., Cleaver, J.E., and Wood, R.D. (1991) *Nucleic Acids Res.* **20**, 991–995.
8. Robins, P., Jones, C.J., Biggerstaff, M., Lindahl, T., and Wood, R.D. (1991) *EMBO J.* **10**, 3913–3921.
9. Guzder, S.N., Sung, P., Prakash, L., and Prakash, S. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 5433–5437.
10. Tomkinson, A.E., Bardwell, A.J., Bardwell, L., Tappe, N.J., and Friedberg, E.C. (1993) *Nature* **362**, 860–862.
11. Tanaka, K., Miura, N., Satokata, I., Miyamoto, I., Yoshida, M.C., Satoh, Y., Kondo, S., Yasui, A., Okayama, H., Okada, Y. (1990) *Nature* **348**, 73–76.
12. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press, Cold Spring Harbor, New York.
13. Sancar, A., Clarke, N.D., Griswold, J., Kennedy, W.J., and Rupp, W.D. (1981) *J. Mol. Biol.* **148**, 63–76.
14. Westerveld, A., Hoeijmakers, J.H.J., Van Duin, M., de Wit, J., Odijk, H., Patsink, A., Wood, R.D., and Bootsma, D. (1984) *Nature* **310**, 425–429.
15. Van Duin, M., de Wit, J., Odijk, H., Westerveld, A., Yasui, A., Koken, M.H.M., Hoeijmakers, J.H.J., and Bootsma, D. (1986) *Cell* **44**, 913–923.
16. Thompson, L.H., Mooney, C.L., and Brookman, K.W. (1985) *Mutation Res.*, **150**, 423–429.
17. Manley, J., Fire, A., Cano, A., Sharp, P., and Gefter, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3855–3859.
18. Huang, J.C., Svoboda, D.L., Reardon, J.T., and Sancar, A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 3664–3668.
19. Nichols, A., and Sancar, A. (1992) *Nucleic Acids Res.* **20**, 2441–2446.
20. Svoboda, D.L., Taylor, J.S., Hearst, J.E., and Sancar, A. (1993) *J. Biol. Chem.* **268**, 1931–1936.
21. Eker, A.P.M., Vermeulen, W., Miura, N., Tanaka, K., Jaspers, N.G.J., Hoeijmakers, J.H.J., and Bootsma, D. (1992) *Mutation Res.* **274**, 211–224.
22. Masutani, C., Sugasawa, K., Asahina, H., Tanaka, K., and Hanaoka, F. (1993) *J. Biol. Chem.* **268**, 9105–9109.
23. Rubin, J.S., Joyner, A.L., Bernstein, A., and Whitmore, G.F. (1983) *Nature* **306**, 206–209.
24. Biggerstaff, M., and Wood, R.D. (1992) *J. Biol. Chem.* **267**, 6879–6885.
25. Sung, P., Prakash, L., and Prakash, S. (1992) *Nature* **355**, 743–745.
26. Bailly, V., Sommers, C.H., Sung, P., Prakash, L., and Prakash, S. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 8273–8277.
27. Bardwell, L., Cooper, A.J., and Friedberg, E.C. (1992) *Mol. Cell. Biol.* **12**, 3041–3049.
28. Davie, J.R., Wynn, R.M., Cox, R.P., and Chuang, D.T. (1992) *J. Biol. Chem.* **267**, 16601–16606.